In proceedings of the *International Conference on Robots and Systems (IROS)*, Pittsburgh PA, August 5-9 1995.

Mobile Robot Localization by Remote Viewing of a Colored Cylinder

Richard Volpe **Todd Litwin** **Larry Matthies**
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

Abstract

To visually determine the position and orientation of a mobile robot from a fixed location in its vicinity, we have employed a cylindrical target which has different colors in each of its four quadrants. By judicious selection of the colors, segmentation of imagery from the fixed location can determine the size and centroid of the cylinder, as well as the visible color quadrants. Both the cylinder size in monocular images, and the centroid disparity in stereo pairs, are shown to provide a measure of distance. The angle of the cylinder is determined by analyzing which color quadrants are visible and to what degree. Implementation and experimental testing of this technique shows that it provides accurate localization data to within one or two pixels of error.

1 Introduction

In 1996, NASA will launch the first of a series of spacecraft to revisit the planet Mars. This *Pathfinder* lander will contain the Microrover Flight Experiment (MFEX), a 10 kg six-wheeled mobile robot which will venture out from the lander, taking pictures and positioning an alpha/proton/x-ray spectrometer against designated soil and rocks [2]. Figure 1 shows our outdoor test system which emulates the Pathfinder scenario.

As MFEX moves about the lander vicinity, its position will be incrementally estimated by wheel encoders and a heading/rate gyroscope [3]. However, the position estimate becomes increasingly erroneous as error accumulates over time [6]. For the flight mission, such errors will be removed by human evaluation of the rover position/orientation from lander stereo imagery obtained daily.

To remove this need for human recalibration and provide accurate rover position estimation throughout day, we have been investigating alternative rover localization schemes. Automatic stereo vision is one obvious extension of the MFEX system. To simplify finding the rover in the imagery, an initial idea of a rover mounted target was introduced. To determine orientation as well, this concept has been extended to a colored cylinder which has a different appearance from all viewing directions. Further, the known size of the cylinder compared with its image size can be used as an alternative method for range determination. This paper describes the geometry of this proposed colored cylinder approach, and presents experimental data verifying the analysis.

Figure 1: *The colored cylinder in field tests sits on top of Rocky 3.2 and is viewed from stereo cameras on top of a tetrahedral Pathfinder lander mockup. The van in the background houses the operator control station.*

Our longterm solution for Mars rover localization will probably involve the replacement of the colored cylinder approach with two more sophisticated methods: position from a radio beacon system, and orientation from measurements of the Sun's position in the sky (combined with knowledge of the time of day). Orientation via compass readings cannot be used since Mars has no appreciable magnetic field [5]. The near-lander behavior of the radio beacon system has yet to be determined, however. One possible augmentation

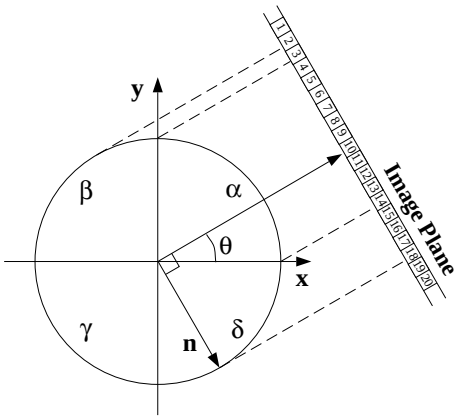

Figure 2: *Geometry of the viewing of the colored cylinder. See the text for details.*

to it may be visual localization of a colored cylinder on the lander from the rover.

Beyond applications for Mars, colored cylinder localization provides a simple and effective method for accurate tracking of mobile robots, both outdoors and in. If the robots are constrained by low mass, power, volume, or cost, on-board sensors for position/orientation may not be feasible. Alternatively, if no stationary beacons or landmarks exist in the work area of the robot, the colored cylinder can be introduced to provide an effective point of reference.

This paper is organized as follows. First, Section 2 reviews the geometry of colored cylinder localization and provides the equations for orientation, size-based range, and stereo-based range. This is followed by an analysis of the sensitivity of the calculations to measurement errors. Section 3 then presents a review of the implementation of this localization method, and the resulting experimental data. Finally, Section 4 describes a possible extension of this technique by utilizing subpixel resolution of cylinder images to obtain more precise results.

2 Rover Visual Localization

2.1 Position and Orientation Determination

Figure 2 shows a top view of a short cylinder of radius $\rho = |\mathbf{n}|$. The external wall of each of the four quadrants is a different color $\kappa \in \{\alpha, \beta, \gamma, \delta\}$. When viewed from a distance, each color will be orthographically projected onto the image plane in quantity $k \in \{A, B, C, D\}$. The image width of the cylinder is m pixels, and the total image width from the camera is M.

2.1.1 Orientation

Initially, it is assumed that the viewing direction is in the first quadrant at an angle θ, and the vector $\mathbf{n}$ is normal to this direction as shown in the figure. Therefore, the amount of each color seen is:

$$A = (\mathbf{n} \cdot \hat{\mathbf{x}}) - (\mathbf{n} \cdot \hat{\mathbf{y}}) \quad (1)$$

$$B = \rho + (\mathbf{n} \cdot \hat{\mathbf{y}}) \quad (2)$$

$$C = 0 \quad (3)$$

$$D = \rho - (\mathbf{n} \cdot \hat{\mathbf{x}}) \quad (4)$$

where the circumflex accent indicates unit vectors. The quantity C is zero since it can not be seen. To remove the dependence on the cylinder radius, it is useful to take the ratio of the quantities: A/B and A/D. Recognizing that

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{x}} = \sin\theta \quad (5)$$

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{y}} = -\cos\theta \quad (6)$$

the ratios may be reformulated as a pair of equations:

$$\begin{bmatrix} B & A+B \\ A+D & D \end{bmatrix} \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} A \\ A \end{bmatrix} \quad (7)$$

Solving this set of equations yields:

$$\sin\theta = \frac{A+B-D}{A+B+D} \quad (8)$$

$$\cos\theta = \frac{A-B+D}{A+B+D} \quad (9)$$

For viewing directions in other quadrants, different colors will be involved and an angular offset is needed (θ_0 equal to a multiple of $\frac{\pi}{2}$). A simple test to determine the quadrant of interest is to measure which color is not present. A zero value for this color quantity specifies the row in the following table:

A	B	C	D	θ_0
k_1	k_2	0	k_3	0
k_3	k_1	k_2	0	$\frac{\pi}{2}$
0	k_3	k_1	k_2	π
k_2	0	k_3	k_1	$\frac{3\pi}{2}$

Using the proper row for the zero color, the angular offset is determined and the measured color quantities are mapped to intermediate variables, k_i, for use in the following orientation solution:

$$\theta = \theta_0 + \text{atan2}\left(\frac{k_1+k_2-k_3}{k_1+k_2+k_3}, \frac{k_1-k_2+k_3}{k_1+k_2+k_3}\right) \quad (10)$$

where $\text{atan2}(x, y)$ is the unambiguous form of $\tan^{-1}(x/y)$.

2.1.2 Position by Size

Figure 3 shows the geometry of viewing at a distance d, with a camera field of view φ. Again assuming an orthographic projection of the scene on to the image plane, the image will consume m/M of the width. Therefore,

$$d = \frac{M\rho}{m \tan\frac{\varphi}{2}} \quad (11)$$

When the camera is close to the cylinder, the perspective can introduce a foreshortening error. Figure 3 shows

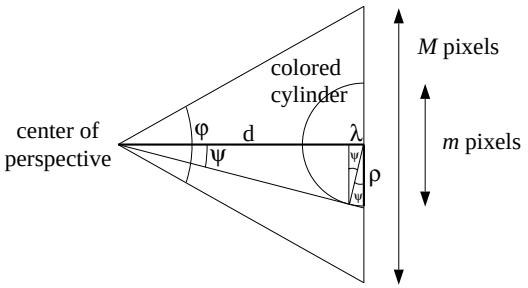

Figure 3: *Geometry for determining the distance to the cylinder. See the text for details.*

that $sin\psi = \frac{\rho}{d} = \frac{\lambda}{\rho}$. Therefore, the correction quantity, λ, must be added to the previous expression for the cylinder distance:

$$d \approx \frac{M\rho}{m \tan \frac{\varphi}{2}} + \frac{\rho^2}{d} \tag{12}$$

In this approximation, Equation (11) can be used to iteratively provide a solution, or Equation (12) can be solved as a quadratic equation.

For example, the 17 pixel image from the previous section corresponds to a distance of 11.24 m, assuming a 0.1 m radius cylinder with a nominal 512 x 480 image and a 30° field of view.

2.1.3 Position by Stereo Image Triangulation

The geometry for stereo triangulation is very similar to that described in the previous section [4]. Instead of relying on the edges of the imaged cylinder to form a triangle with the camera location, two images are used to triangulate on the centroid of the cylinder. In this case, the distance is dependent on the *disparity* of the centroid in the images, measured in pixels μ:

$$d = \frac{Mb}{2\mu \tan \frac{\varphi}{2}} \tag{13}$$

where b is the baseline distance between the stereo cameras. Since the cylinder is radially symmetric, no view angle effects are encountered to influence the centroid location.

2.2 Pixel Resolution Effects

2.2.1 Orientation Errors

The precision of this orientation measuring scheme is dependent on the spatial resolution of the pixels in the image of the cylinder. There are two extremes to the resolution of orientation determination by this scheme, depending on the amount of rotation needed to move the color boundaries from one pixel to the next.

The *best case* of orientation resolution occurs when one of the color boundaries is in the center of the image. In this orientation, a minimal amount of rotation is needed, since the motion of the color boundary is approximately parallel to the image plane. In the general case of m total pixels, the angular resolution per pixel is:

$$\Delta\theta_{min} = \frac{max\ image\ plane\ motion}{radius\ of\ cylinder} \tag{14}$$

$$= \left(\frac{2\rho}{m}\right)\left(\frac{1}{\rho}\right) = \frac{2}{m} \tag{15}$$

The *worst case* of orientation resolution occurs when the color boundary motion is most out of the image plane. Given the four quadrant design of the cylinder, this occurs for $\theta \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$. Therefore,

$$\Delta\theta_{max} = \left(\frac{2\rho}{m \cos \frac{\pi}{4}}\right)\left(\frac{1}{\rho}\right) = \frac{2\sqrt{2}}{m} \tag{16}$$

For the example of $m = 17$, the resultant resolution bounds are: $\Delta\theta_{min} = 6.7°$ and $\Delta\theta_{max} = 9.5°$.

It is useful to eliminate m from the equations above, by employing Equation (11) or (12):

$$\Delta\theta_{min} = \frac{2d \tan \frac{\varphi}{2}}{M\rho} \tag{17}$$

$$\Delta\theta_{max} = \sqrt{2}\Delta\theta_{min} \tag{18}$$

2.2.2 Relative Position Error by Size

For distance measurements, *relative* error is used to normalize the results and allow more direct comparison of them. The relative error in the distance calculation of Equation (11) caused by a one pixel change in the size is:

$$\left.\frac{\Delta d}{d}\right|_{size} = -\frac{M\rho}{dm^2 \tan \frac{\varphi}{2}} = -\frac{1}{m} = -\frac{d \tan \frac{\varphi}{2}}{M\rho} \tag{19}$$

In the example, a one pixel change causes an error of 0.66 m, or about 5.8% of the total distance.

2.2.3 Relative Position Error by Stereo

Similarly, using Equation (13) the relative error for distance by stereo caused by a one pixel change in the centroid is:

$$\left.\frac{\Delta d}{d}\right|_{stereo} = -\frac{Mb}{2d\mu^2 \tan \frac{\varphi}{2}} = -\frac{1}{\mu} = -\frac{2d \tan \frac{\varphi}{2}}{Mb}. \tag{20}$$

If the stereo camera baseline is equal to the diameter of the cylinder, the resolution will be same as in the last section.

3 Implementation and Experimentation

3.1 Color Segmentation Method

To experimentally determine the accuracy of the this localization scheme, a cylinder was constructed out of fluorescent colored cardboard. The cylinder is 19.2 cm in diameter by 6.4 cm high, and the four quadrants are colored blue, magenta, red, and magenta again. The colors were originally selected so that adjacent colors have similar spectral

response in one channel of the color cameras, helping prevent boundary pixels from being excluded in the segmentation of the cylinder from the full image.

The segmentation of the image to locate the quadrants of the cylinder is performed in four steps: (1) the image pixels are classified by color, (2) nearly adjacent regions are connected, (3) the cylinder is located, and (4) the visible colored quadrants of the cylinder are located.

Three color classes are defined, one for each of the three distinct colors of the cylinder's quadrants: blue, magenta, and red. Determining to which class each pixel belongs is done by taking the ratio of the primary colors of the image (for the red and magenta classes: green/red and blue/red; for the blue class: red/blue and green/blue). For each ratio an experimentally determined range of values is used to determine class membership. Intensity ratios are used because of their relative insensitivity to ambient lighting.

As mentioned previously the choice of colors is made to minimize rejection of pixels on the transition between the colors. Even so, many pixels on such boundaries are rejected as belonging to neither color class. As will become clear in the next step, it is important that the three visible color quadrants of the cylinder need to touch each other in the image. It is therefore necessary to filter the image in order to join almost touching groups of pixels of different colors. This is done by finding any pixel separated on its right by only background (rejected) pixels from a pixel of a different color. The gap considered is from one to four background pixels. The gap is filled in symmetrically by growing the colors on the left and the right of the gap toward each other.

Before searching for the cylinder in the image, the area of interest in the image is restricted so that none of the sky is visible when operating outdoors. Otherwise, the sky is often confused with one of the quadrant colors. Region growing is then performed on the selected part of the image using local blob coloring to identify the four-connected regions of the image which contained any mixture of the cylinder's colors [1]. Without any analysis as to shape, the largest such region is taken as corresponding to the cylinder. This is not unreasonable since the colors were quite artificial in appearance, and the background is typically a natural outdoors scene. Also, in a later error-checking stage, the bounding boxes of the cylinder in the left and right stereo images are compared to make sure that they are similar to each other.

The bounding box of the cylinder is then analyzed to find the individual quadrants. The same region-growing technique used above is again applied here several times more. But this time, instead of growing regions of any mixture of the colors, only regions of single colors are considered. The largest such region is labeled the "middle" region. Looking to the left of the middle region, the largest region of a different color is labeled the "left" region. Similarly the largest region to the right of the middle is labeled the "right" region. Consistency checking is performed to make sure that the regions found make physical sense. Note that it is possible for only one or two regions to be visible.

The final result of this process is the center of mass of the entire cylinder, as well as the bounding boxes and pixel counts for each of the quadrants.

3.2 Summary of the Experimental Results

Tests of the cylinder localization were performed with its center along the line midway between a pair of Sony XC999 cameras, set up for stereo viewing with a baseline of 12 cm. Calibration procedures indicate that the field of view of the cameras is 29.33°.

Measurements of the cylinder were conducted at 20° increments and five different distances. The zero position of the cylinder was aligned with the center of rotation of the cameras by bore-sighting, and the cylinder was rotated with a machinists' turntable. This procedure was repeated at each of the chosen distances of 0.914, 1.829, 4.572, 10.058, and 15.240 meters (3, 6, 15, 33, and 50 feet, respectively) from the center of rotation of the stereo camera fixture. The front of the cameras was offset from this origin by .076 m (3 inches). Ten measurements were taken at each position/orientation setting.

Figures 4–6 show the mean measured angular error, and the mean measured relative distance error for size and stereo. All data is plotted as a function of true cylinder angle. It can be seen that the angular errors are typically within ±5°, especially when the cylinder is close. For size-based distance determination, the means of the relative errors are off by as much as 15% for the 10 meter data. Worse still, the 15 m data is very inaccurate. However, for stereo distance determination the results are much better. The calculated distance is typically within 5% of correct, and contains a constant bias which can be eliminated by calibration.

3.3 Discussion of the Experimental Results

It is apparent from the results that within a 10 meter range, the colored cylinder localization technique provides reasonably accurate position and heading information for the rover: typically within 5° of true heading and 5% of true distance.

The standard deviations shown in Figures 7–9 further indicate that this measurement technique is quite precise, especially at ranges less than 10 m. These graphed results should be compared against the calculated errors for one pixel changes in the cylinder images, provided in the following table.

d	$\Delta\theta_{min}$	$\Delta\theta_{max}$	$\frac{\Delta d}{d}\|_{size}$	$\frac{\Delta d}{d}\|_{stereo}$
0.838 m	0.50°	0.71°	0.0045	0.0071
1.753 m	1.06°	1.50°	0.0093	0.015
4.496 m	2.70°	3.82°	0.024	0.038
9.982 m	6.09°	8.61°	0.053	0.085
15.164 m	9.25°	13.43°	0.081	0.129

It is apparent that for both orientation and distance, the standard deviations indicate that the segmented images usually deviate by less than one or two pixels. Therefore, once calibrated for bias, this method will provide results that are at the limit of measurement of the given cameras.

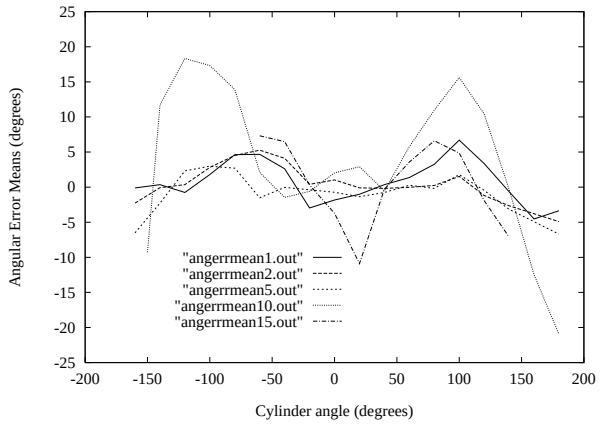

Figure 4: *Means of the angular error at the five tested distances.*

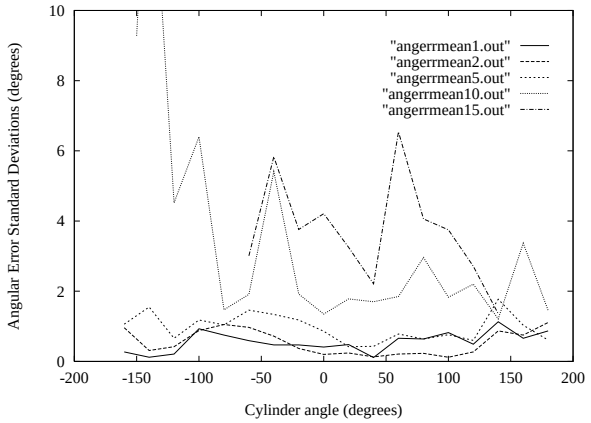

Figure 7: *Standard deviation of the angular error at the five tested distances. The one data point off the graph is at (-140,16).*

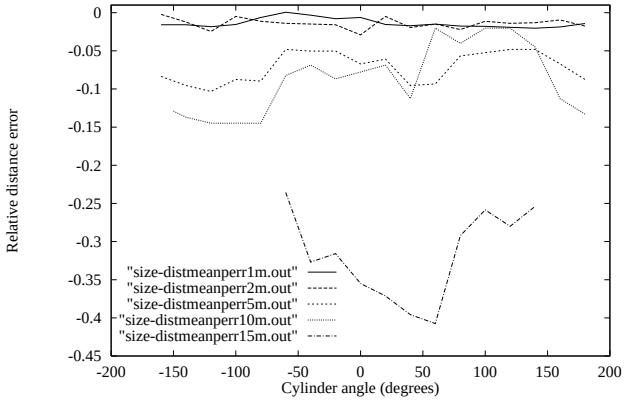

Figure 5: *Means of the relative error for size distance determination at the five tested distances.*

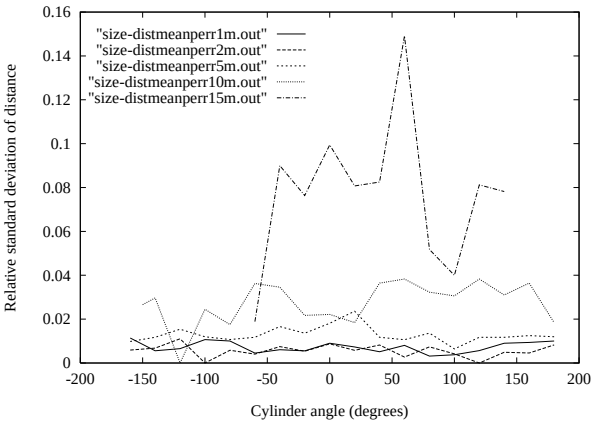

Figure 8: *Relative standard deviations for size distance determination at the five tested distances.*

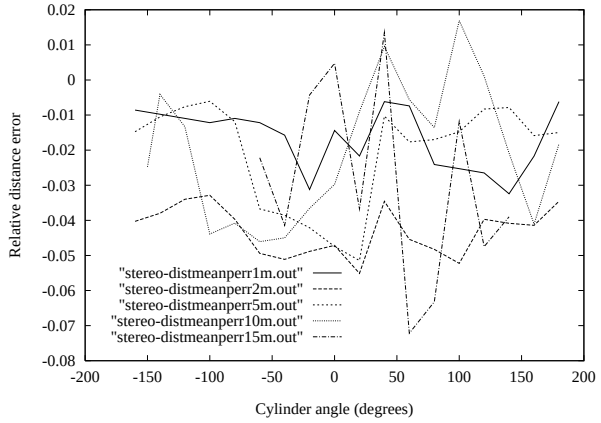

Figure 6: *Means of the relative error for stereo distance determination at the five tested distances.*

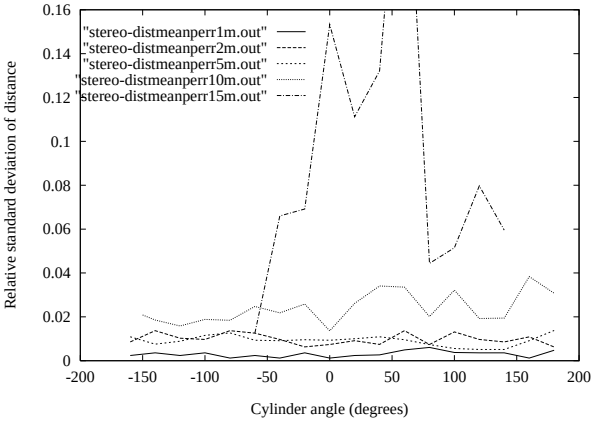

Figure 9: *Relative standard deviations for stereo distance determination at the five tested distances. The point off the plot is at (60,0.26).*

Of notable exception to the generally good quality of measurements is the angular error and standard deviation in the $-180 \leq \theta \leq -90$ range for the 10 m data. This is attributed to problems in segmenting the magenta and red quadrants at this distance. Subsequent experiments at the 15 meter range further confirmed this conclusion by providing reasonable results except in this angular domain, where no consistent measurements were possible.

Two improvements to the implementation of this technique would be a selection of different colors and use of a more robust segmentation algorithm. One possibility for new colors would be a simple replacement of the red panel with a green one. A more strongly contrasting color appears to be desirable since the contrast between the magenta and blue enabled more accurate cylinder orientation determination. This can be seen in the 10 meter angular error data — the smallest errors were in the range of $-90°$ to $+90°$, when the blue quadrant plays a significant role. (Partial data taken at 15 meters confirmed this trend.)

The second improvement, robust segmentation, would increase the accuracy of both the angle and distance measurements. First, by refining the segmentation of the individual colored regions on the cylinder, the angle measurements would improve. Second, better segmentation of the whole cylinder would improve size based distance determination. Third, by fine tuning the the accuracy of the cylinder centroid measurement, stereo disparity calculations and the resultant distance determination would improve slightly. Finally, episodes of erroneous segmentation of regions of the image that are not part of the cylinder could be reduced by an improved algorithm.

4 Extension to Subpixel Resolution

To further improve these results, it would be necessary to determine color quadrant image boundaries within the boundary pixels. For instance, in Figure 2 subpixel resolution is needed in pixels 2, 3, 14, and 18.

The possibility of determining the subpixel color boundary is dependent on the camera technology employed. For instance, a camera like the XC999 used in the previously discussed tests has a single color CCD chip with each of its pixels divided into four color elements (e.g. yellow, cyan, magenta, and green). Unfortunately, this technology does not enable subpixel resolution for our purposes. Other cameras, however, employ three separate CCDs, each covering the same line of sight but sensitive to different colors (e.g. red, green, and blue). With this second type of camera, or with a black and white camera and three colored filters, subpixel resolution is directly achievable.

Consider the situation shown in Figure 10, where the image of the boundary between two colored regions falls in the middle of a pixel. In this case, the color sensed by the pixel will be the weighted average of the two colors:

$$\kappa = w_{ij}\kappa_i + (1 - w_{ij})\kappa_j \tag{21}$$

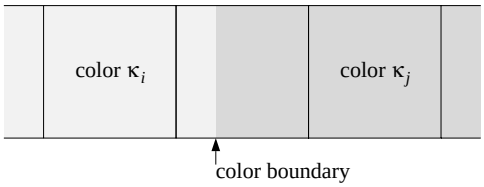

Figure 10: *Geometry for subpixel resolution. See the text for details.*

or

$$w_{ij} = \frac{\kappa - \kappa_j}{\kappa_j - \kappa_i} \tag{22}$$

where w_{ij} is the percentage of area in the boundary pixel covered by the first of the two colors, κ_i and κ_j. The sensed color quantities may, therefore, be refined as:

$$k_i = k_i' + \sum_{j \neq i} w_{ij} \sum_{\ell}^{4} \frac{k_\ell}{m} \tag{23}$$

where k_i' is the color quantity obtained from only those pixels directly matching the i^{th} cylinder color. The summation over ℓ is necessary for normalization. This analysis is not only applicable on the boundaries between cylinder colors, but also between the cylinder and the background. In this case, only the background/cylinder border weights are needed.

The subpixel calculations have an inherent resolution as well. This is due to the spectral resolution of the camera, 2^n, where n is the number of bits per pixel. Since the color change of the boundary pixel is a linear function of the boundary position, the discretization of the pixel color gives the same resolution spatially. Therefore, the angular and radial resolution in Equations (15), (16), (19), and (20) is reduced by the factor 2^n. For an 8-bit image the new values for the resolution of our example are: $0.26° \leq \Delta\theta \leq 0.37°$, $\Delta d = 2.6\text{mm} = 0.023\%$ of d. Even allowing that the full 8-bit color signal may not be available due to noise and other factors, considerable improvement in localization measurements can be expected with subpixel calculations.

5 Conclusion

We have presented a simple method for determining the position and orientation of a mobile robot by viewing a colored cylinder placed upon it. Experiments have shown that this method provides results accurate to one or two pixels. Small errors, therefore, require images at least an order of magnitude larger ($\sim$20 pixels). For our experimental setup, this translated into an effective range of about 10 m.

It important to point out that these results are pertinent to the Pathfinder MFEX. Although human orientation determination of the rover will be used, the visual clues will be limited by similar image resolution problems. Therefore, rover operations will need to be constrained to distances that provide images a few tens of pixels in width. The Pathfinder lander camera specifications provide a similar range limit of $\sim$10 m.

6 Acknowledgments

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

References

[1] D. Ballard and C. Brown. *Computer Vision.* Prentice-Hall, Englewood Cliffs, NJ, 1982.

[2] T. Economou et al. Use of Mercuric Iodine X-Ray Detectors with Alpha BackScattering Spectrometers for Space Applications . *IEEE Transactions on Nuclear Science*, 38(2):574–579, 1991.

[3] E. Gat et al. Behavior Control for Robotic Exploration of Planetary Surfaces. *IEEE Transactions on Robotics and Automation*, 10(4):490–503, 1994.

[4] B. Horn. *Robot Vision.* MIT Press, Cambridge, MA, 1986.

[5] H. Kieffer (editor). *Mars.* University of Arizona Press, 1992.

[6] L. Matthies et al. Mars Microrover Navigation: Performance Evaluation and Enhancement. In *submitted to 1995 Internation Conference on Robots and Systems* , Pittsburgh, PA, August 5-9 1995.